

AD-A122 773

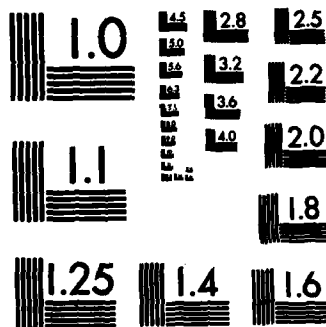
DESIGN ISSUES FOR MIXED MEDIA PACKET SWITCHING NETWORKS
(U) HAWAII UNIV HONOLULU D D HUYNH ET AL. DEC 75
B75-38 N00014-76-C-0256

1/1

UNCLASSIFIED

F/G 17/2 NL

								END					
								FILED					
								1 DEC					



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

DESIGN ISSUES FOR MIXED MEDIA PACKET SWITCHING NETWORKS

BY

D.D. HUYNH, H. KOBAYASHI, AND F.F. KUO

AL A12273

DTIC FILE COPY

DECEMBER 1975

DTIC
ELECTE
DEC 29 1983
S H D

APPROVED FOR PUBLIC RELEASE
DISTRIBUTION UNLIMITED

TECHNICAL REPORT B75-30

D

DESIGN ISSUES FOR MIXED MEDIA PACKET SWITCHING NETWORKS

by

D.D. Huynh, H. Kobayashi and F.F. Kuo
THE ALOHA SYSTEM
University of Hawaii
Honolulu, Hawaii

DTIC
ELECTE
DEC 29 1983
S D
H

December 1975

Sponsored by

Advanced Research Projects Agency*
ARPA Order No. 1956
and
ONR Contract No. N00014-76-C-0256

DISTRIBUTION STATEMENT A
Approved for public release
Distribution Unlimited

* The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the Advanced Research Projects Agency of the United States Government.

DESIGN ISSUES FOR MIXED MEDIA PACKET SWITCHING NETWORKS

by

D. Huynh, H. Kobayashi* and F.F. Kuo**
THE ALOHA SYSTEM
University of Hawaii
Honolulu, Hawaii

Abstract

↓

In this paper we present some of the important design issues for packet switching networks with both satellite and terrestrial components -- which we call *mixed media* packet switching networks. Satellite packet switching has considerable promise for low cost, high bandwidth data communications. However there is inherent high delay in satellite links which do not appear in ground links. Therefore a mix of the two communications media offers the advantages of low cost/high bandwidth together with low delay communications where required. In this paper we examine a number of tradeoffs which offer guidelines for the design and optimum utilization of mixed media networks.

DTIC
COPY
INSPECTED
3

Accession For	NTIS GRA&I	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	DTIC TAB	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Unannounced	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Justification	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
By	Distribution/			
Availability Codes				
Dist	Avail and/or Special			
		A		

This report was supported by THE ALOHA SYSTEM, a research project at the University of Hawaii, which is supported by the Advanced Research Projects Agency of the Department of Defense and monitored by NASA Ames Research Center under Contract No. NAS2-8590 and by ONR Contract No. N00014-76-C-0256.

* Dr. Kobayashi is on leave from IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598.

** Dr. Kuo is now with the Department of Defense, Washington, D.C., 20301.

Design Issues for Mixed Media Packet Switching Networks

by D. HUYNH, H. KOBAYASHI* and F. F. KUO**

THE ALOHA SYSTEM

University of Hawaii

Honolulu, Hawaii

INTRODUCTION

In 1968 the Advanced Research Projects Agency (ARPA) of the U. S. Department of Defense began implementation of a computer-communication network which permits the interconnection of heterogeneous computers at geographically distributed centers throughout the United States. This network has come to be known as the ARPANET [1,2], and has grown from the initial four node configuration in 1969 to almost forty nodes (including satellite nodes in Hawaii, Norway, and London) in late 1974. The major goal of ARPANET is to achieve resource sharing among the network users. The resources to be shared include not only programs, but also unique facilities such as the powerful ILLIAC IV computer and large global weather data bases that are economically feasible when widely shared.

The ARPANET employs a distributed store-and-forward packet-switching approach that is much better suited for computer-communication networks than the more conventional circuit-switching approach. Reasons favoring packet switching include lower cost, higher capacity, greater reliability and minimal delay. The CCITT (Comité Consultatif International Télégraphique et Télé-

This report was supported by THE ALOHA SYSTEM, a research project at the University of Hawaii, which is supported by the Advanced Research Projects Agency of the Department of Defense and monitored by NASA Ames Research Center under Contract No. NAS2-8590 and by ONR Contract No. N00014-76-C-0256.

* Dr. Kobayashi is on leave from IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598.

**Dr. Kuo is now with the Department of Defense, Washington, D.C., 20301.

phonique), an international standards organization, defines a packet as "a group of binary digits including data and call control signals which is switched as a composite whole. The data, call control signals and possibly error control information are arranged in a specific format."

Multi-access satellite channels

Communication satellites operating in packet mode are becoming increasingly important for consideration in the design of large computer communication networks. In particular, most possess the following characteristics which are of special importance: the satellite's antenna coverage allows any of a large number of ground stations to access it at any time (multi-access), and its transmissions can be received by all of these stations at all times (broadcast).

Up to the present, packet switched computer networks have mainly utilized terrestrial communications links. Recently, Telenet Communications Corporation, one of the new value-added carriers [3] announced plans to offer public packet switched data service in which terrestrial and satellite links will be available. ARPA has also been planning to augment its terrestrial links with satellite communications and has commissioned several satellite IMPs or SIMPs [4] to be built by Bolt Beranek and Newman Inc. (BBN). In mid 1975, BBN and the British Post Office jointly conducted a multi-access satellite experiment using the Atlantic INTELSAT IV satellite with SIMPs located at the ETAM ground station in the U. S. and the Goonhilly Downs ground station in England.

The basis of this experiment is the work on the ALOHANET of THE ALOHA SYSTEM project at the University of Hawaii [5,6,7]. The multiplexing technique that is used by the ALOHANET is a multi-access packet switching method that has come to be known as the ALOHA technique.

Based upon the ALOHA multiplexing method or variations thereof, a number of techniques have been proposed for the utilization of a satellite channel in a packet-switched data network in a way which allows all stations to dynamically share the channel capacity [8,9,10,11,12,13,14]. All are based on the division of the channel into time slots approximately equal to a packet transmission time. The approach described by Roberts [8] and Kleinrock and Lam in [12,13] makes use of a technique called "slotted ALOHA random access" or "slotted ALOHA". The channel capacity of a slotted ALOHA channel was estimated to be $1/e \approx 36\%$ [8].

Goals of this paper

In this paper we will examine a number of key issues in the design of a packet-switched communication network composed of a terrestrial store-and-forward packet switching component combined with a multi-access/broadcast satellite as depicted in Figure 1.

The factors involved in the design optimization of a packet switched network are [15]:

- Node location and traffic matrix
- Topology of links
- Capacity of links
- Routing
- Flow control, other network design, etc.

To attempt to optimize any one factor, such as the capacity of links, it is necessary to assume that the factors higher in the list must be given. However, to decide optimum choices for the factors high on the list may require detailed calculations of all the lower ones. Fortunately, such factors as

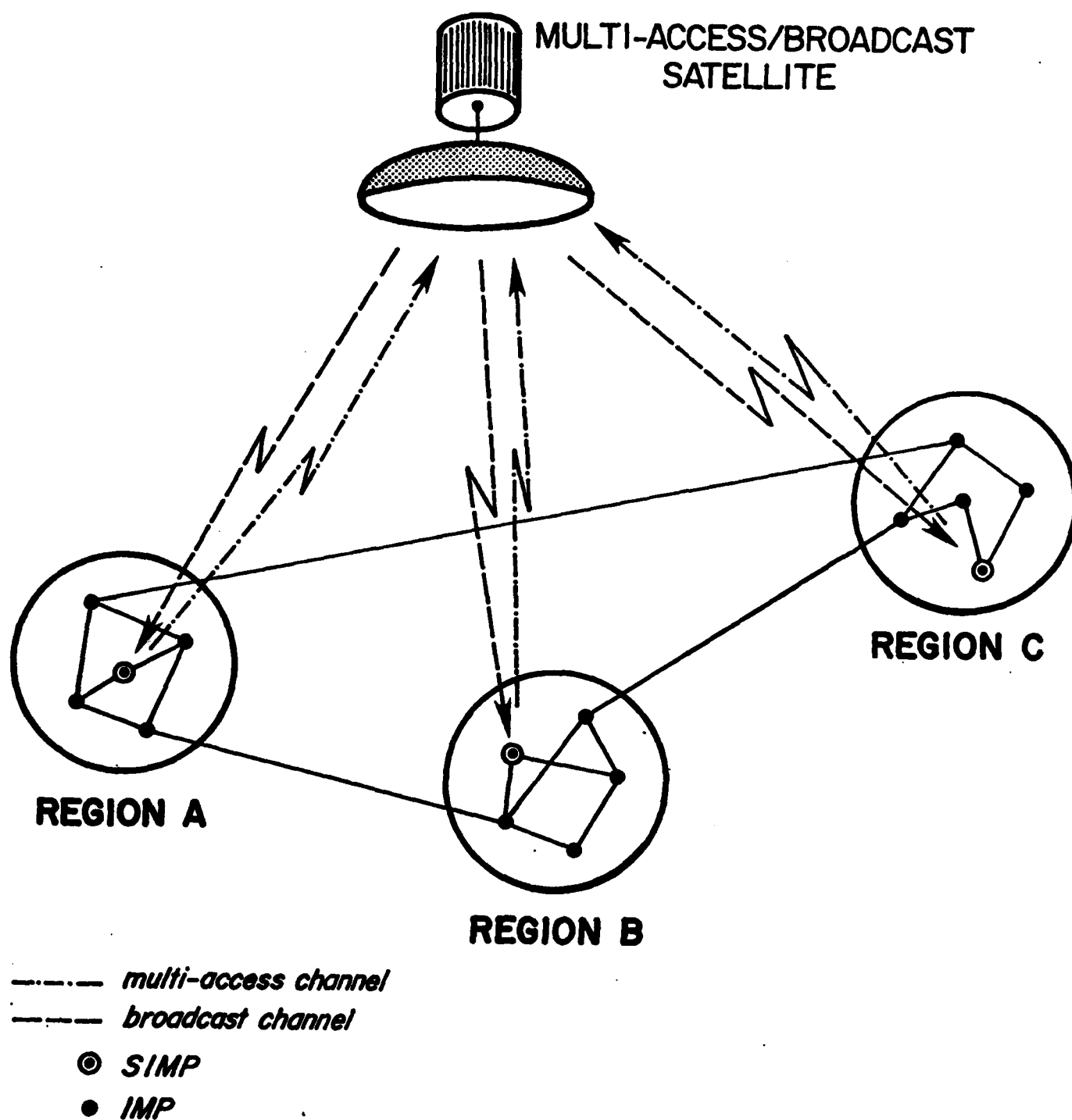


Figure 1
THE PROPOSED NETWORK MODEL

node location are usually not design variables, being fixed by practical considerations. Frank, Frisch and Chou [16] have developed suboptimum procedures for the topological layout of a packet switched network. Thus we will concentrate on the following problems in our paper.

1. Routing of packets via ground or satellite
2. Capacity assignments for ground and satellite channels
3. Retransmission strategies

In this paper we will concentrate on the discussion of design issues such as throughput, delay, cost, etc., rather than dwell on the theoretical development of the design equations which is given in a companion paper to be published elsewhere [17].

THE NETWORK MODEL

The network model under consideration consists of a terrestrial store-and-forward packet switching network referred to here as the ground subnet, and a multi-access/broadcast satellite which together with the SIMP ground stations is the satellite subnet.

For the ground subnet we will use the model given in the paper by Crowther, et. al. [18], and will reproduce the succinct definitions of network terminology given in that paper:

Nodes. The nodes of the network are real-time computers, with limited storage and processing resources, which perform the basic packet-switching functions.

Hosts. The Hosts of the network are the computers, connected to nodes, which are the providers and users of the network services.

Lines. The lines of the network are some type of communications circuit of relatively high bandwidth and reasonably low error rate.

Connectivity. We assume a general, distributed topology in which each node can have multiple paths to other nodes, but not necessarily to all other nodes. Simple networks such as stars or rings are degenerate cases of the general topology we consider.

Message. The unit of data exchanged between source Host and destination Host.

Packet. The unit of data exchanged between adjacent nodes.

Acknowledgement. A piece of control information returned to a source to indicate successful receipt of a packet or message. A packet acknowledgement may be returned from an adjacent node to indicate successful receipt of a packet; a message acknowledgement may be returned from the destination to the source to indicate successful receipt of a message.

Store and Forward Subnetwork. The node stores a copy of a packet when it receives one, forwards it to an adjacent node, and discards its copy only on receipt of an acknowledgement from the adjacent node, a total storage interval of much less than a second.

Packet Switching. The nodes forward packets from many sources to many destinations along the same line, multiplexing the use of the line at a high rate.

Routing Algorithm. The procedure which the nodes use to determine which of the several possible paths through the network will be taken by a packet.

For the satellite subnet we use the finite population model of the slotted ALOHA channel developed by Kleinrock and Lam [13], which we call *Scheme 1* and we also consider a multi-access channel model in which there is no retransmission via satellite. This model which we denote as *MASTER* or *Scheme 2* is described in a later section of this paper.

The combination of a system that operates in contention mode (the satellite) and one that operates in queueing mode (the terrestrial store-and-forward net) into an overall system model presents many interesting problems. These problems are mainly due to the fact that with contention systems the throughput increases to a maximum and then decreases as system load increases while for queueing systems the throughput increases to one as system load increases [13]. The model of a multi-access slotted-ALOHA channel is typical of a contention model, and to use the channel optimally the system load on the channel must be carefully controlled. We have tried to develop an analytical model of the overall system so that we may predict and optimize its

performance. Our network model thus consists of:

- A. A set of store-and-forward IMP-like devices interconnected by capacity limited ground channels (a distributed subnet). For the sake of reliability this subnet is at least 2-connected.
- B. A set of SIMP-like devices directly connected to satellite ground stations. These SIMPs are usually geographically scattered and relatively far apart from each other.
- C. A multi-access/broadcast satellite transponder linking all SIMPs in a star configuration.

A SIMP is usually co-located with an IMP at a node and both devices have buffering and scheduling capabilities. We assume that the ground net is regionalized. That is, the ground net is partitioned into regions. Every region has a SIMP and the IMPs in that particular region can only go through one and only one SIMP. The regionalization is determined by the closeness of an IMP to a SIMP in terms of the number of hops and the distance between them. Such a structure is shown in Figure 1.

ROUTING

With a mixed media network the issue of routing is a major concern. With two possible courses to choose from--one via satellite and one via ground, the issue is to choose the set of routes so as to minimize the overall average network delay. The tradeoffs to consider are these; the satellite channel has an inherent minimum delay of .26 seconds. However, satellite capacity is less costly than ground channel capacity for medium to long distances, and therefore more satellite capacity is available at less cost than comparable facilities on the ground. The ground channels are inherently faster than the satellite channel, but because of capacity limitations, are subject to queueing delays, which combined with the store-and-forward nodal processing delays may, in heavy traffic situations, result in larger overall delays than the satellite delays. To summarize, satellite channels have greater delays but also more cost effective channel bandwidth than ground channels.

The routing procedure used in the ARPANET is a distributed adaptive algorithm in which each node has a routing table which is periodically updated with minimum distance estimates from its immediate neighbors [19]. Unfortunately the distributed adaptive algorithm is extremely difficult to describe analytically and could not be applied to our analytical model. Instead we have chosen a deterministic split traffic (bifurcated) routing strategy [19] which because it allows traffic to flow on more than one path between a given source-destination node pair gives a better balance than a fixed routing procedure. It should be noted that our analytical model does not require any specific routing algorithm, but can accommodate any routing algorithm that can be modelled mathematically.

Optimal routing of packets

Using the network model of Figure 1 the main problem in packet routing is to route packets from one region to another via either ground or satellite links in such a way as to minimize the overall average delay of the network. Given the topology of the network and the capacities of the ground and satellite links, we assume a demand matrix $[\gamma_{ij}]$ where γ_{ij} is the average Poisson input in packets/sec from node i to node j . Let us define the routing index g_{ij} as the fraction of γ_{ij} sent through the ground net and let $\bar{g}_{ij} = 1 - g_{ij}$ be the fraction of γ_{ij} sent through the satellite net by first routing to the SIMP in the region where IMP i resides, then transmitting through the satellite channel to the SIMP in the region where IMP j resides. As a result, $g_{ij} = 1$ if both IMPs i and j are in the same region.

Let us define T as being the overall average delay of the network, averaged over all source-destination node pairs and all routes dictated by the routing algorithm. The optimal routing strategy for mixed media networks can be stated as:

$$\min_{g_{ij}} T; \text{ subject to } \{0 \leq g_{ij} \leq 1 \text{ for all } i, j\}$$

For our analysis, we define the following notation:

$$\gamma = \sum_{i,j} \gamma_{ij} \quad \text{total traffic in the net}$$

$$\tau_{ij} = \text{average delay for a packet travelling from IMP } i \text{ to IMP } j$$

$$\sigma(i) = \text{index of the region (hence that of the corresponding SIMP) to which IMP } i \text{ belongs}$$

τ_s = average delay for a packet transmitted from one SIMP to another via the satellite channel

T_{ij} = overall average delay for a packet travelling from IMP i to IMP j (averaged over ground and satellite links)

We can readily derive the following relationships:

For the ALOHA scheme

$$T_{ij} = g_{ij}\tau_{ij} + \bar{g}_{ij}(\tau_{i\sigma(i)} + \tau_s + \tau_{\sigma(j)j}) \quad (1)$$

and

$$T = \frac{1}{Y} \sum_{i,j} Y_{ij} T_{ij}$$

$$= \frac{1}{Y} \sum_{i,j} Y_{ij} [g_{ij}\tau_{ij} + \bar{g}_{ij}(\tau_{i\sigma(i)} + \tau_{\sigma(j)j})] + \frac{1}{Y} \sum_{i,j} \bar{g}_{ij} Y_{ij} \tau_s \quad (2)$$

For the MASTER scheme, in order to obtain an expression analogous to Equations (1) and (2), we further introduce the notation:

P_σ = probability that a packet transmitted from SIMP σ into the satellite channel is successful

Then we obtain

$$T_{ij} = g_{ij}\tau_{ij} + \bar{g}_{ij}[\tau_{i\sigma(i)} + \tau_s + P_{\sigma(i)}\tau_{\sigma(j)j} + (1-P_{\sigma(i)})\tau_{\sigma(i)j}] \quad (3)$$

and

$$T = \frac{1}{Y} \sum_{i,j} Y_{ij} T_{ij}$$

$$= \frac{1}{Y} \sum_{i,j} Y_{ij} \{g_{ij}\tau_{ij} + \bar{g}_{ij}[\tau_{i\sigma(i)} + P_{\sigma(i)}\tau_{\sigma(j)j} + (1-P_{\sigma(i)})\tau_{\sigma(i)j}]\}$$

$$+ \frac{1}{Y} \sum_{i,j} \bar{g}_{ij} Y_{ij} \tau_s \quad (4)$$

In Equations (2) and (4) the terms with g_{ij} represent the portion of traffic from IMP i to IMP j routed entirely through the ground net and those terms with \bar{g}_{ij} represent the portion routed from node i to the regional SIMP $\sigma(i)$, from $\sigma(i)$ to $\sigma(j)$, the regional SIMP in which IMP j resides and then from $\sigma(j)$ to IMP j . Thus even in the satellite routing, there is some associated ground traffic to get to and from the regional SIMPs. Our mathematical model is flexible enough to permit SIMPs at every node, and if this were the case, obviously the $\tau_{i\sigma(i)}$ and $\tau_{\sigma(j)j}$ terms would be zero. The term τ_s is derived from the finite population model of a slotted ALOHA channel as originally given in Lam [13] and in a slightly different form by the authors [17].

Let λ_ℓ be the traffic on link ℓ . We can derive λ_ℓ as a function of g_{ij} , the traffic demand γ_{ij} and the routing algorithm. Since the equation is fairly complex we will not give it here but refer the reader to a companion paper for its derivation [17]. Given λ_ℓ and neglecting the delay due to transmission errors and nodal processing and ground propagation time, the average delay along any ground link can be described by the M/M/1 queueing model by virtue of the well-known independence assumption and Jackson's decomposition theorem [22]

$$T_\ell = \frac{1}{\mu_\ell C_\ell - \lambda_\ell} \quad (5)$$

where C_ℓ is the capacity of channel ℓ and $\frac{1}{\mu_\ell}$ the average packet length on that channel. The overall average delay can be expressed in terms of the link traffic λ_s , $\{\lambda_\ell\}$ and link delays τ_s , $\{T_\ell\}$ as

$$T = \frac{1}{\gamma} \left[\sum_{\ell=1}^L \lambda_\ell T_\ell + \lambda_s \tau_s \right] \quad (6)$$

where L = total number of links in the net, and λ_s the total traffic rate

on the satellite channel, i.e.,

$$\lambda_s = \sum_i \sum_j \bar{g}_{ij} \gamma_{ij}$$

We can show that T in Equation (6) is convex with respect to g_{ij} so that we can obtain a minimum overall delay by varying g_{ij} according to a selected optimization algorithm. The optimum choice of g_{ij} depends also upon the topology of the network, the location of the SIMPs, the capacities of the ground and satellite links, the specific terrestrial routing algorithm, as well as the traffic demand matrix $[\gamma_{ij}]$, all of which must be specified in advance of the optimization computation.

Since the optimum selection of g_{ij} is not critically dependent upon a particular optimization procedure, we used the algorithm known as the BOX COMPLEX procedure [20] since this is one of the optimization algorithms which does not require derivatives of the objective function.

As an example, consider the network model in Figure 2 in which the two regions consist of nodes $\{1,2,3,4\}$ and nodes $\{5,6,7,8\}$, and the regional SIMPs are located at nodes 1 and 7. The traffic demand matrix assumed to be uniform with $\gamma_{ij} = 20$ packets/sec. $i \neq j$ and $\gamma_{ij} = 0$ for $i = j$. The average packet lengths is assumed to be 512 bits on all ground channels. The packet length on the satellite channel is fixed and equals 1 K bit. The ground link capacities are all assumed to be $C_g = 50$ K bits/sec and the satellite capacity to be $C_s = 1500$ K bits/sec. With this information as input to the routing optimization program, the inter-regional g_{ij} are computed and are given in Table 1. Note that we assume $g_{ij} = 1$, for i, j in the same region.

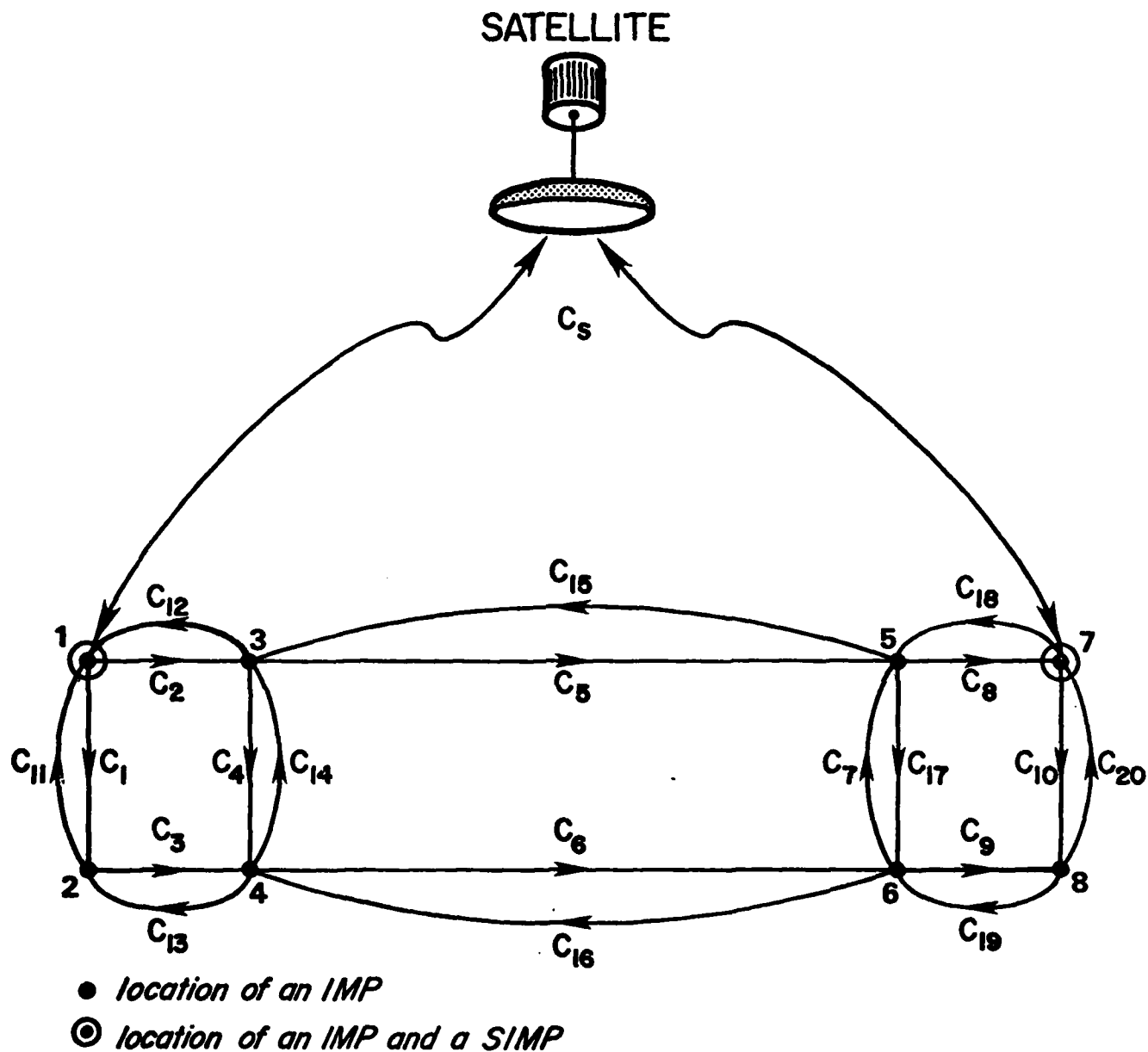


Figure 2

**THE NETWORK MODEL
 CONSIDERED IN NUMERICAL EXAMPLES**

	<i>destination</i>			
	5	6	7	8
1	1	0	0	0
<i>origin</i> 2	.855	.681	0	.706
3	1	1	0	.315
4	1	1	0	1

Table 1

THE ROUTING INDICES FOR TRAFFIC FROM REGION 1 TO REGION 2

Note: The Routing Indices for traffic from Region 2 to Region 1 are symmetrical to those given in Table 1. For instance, $g_{82}=g_{28}=.706$; $g_{72}=g_{18}=0$; etc.

LINK CAPACITY ASSIGNMENTS

In the packet routing studies above we assume that the ground link capacities are given. We also assume as given the capacity of the satellite channel. The capacity assignment problem is: for each channel ℓ the capacity C_ℓ must be found which minimizes the total average message delay T . The problem is most difficult if the capacities must be chosen from a discrete set of options. Following Kleinrock [21] we assume that the capacities of ground links and satellite channel are continuous variables and use analytic procedures involving Lagrange multipliers to obtain C_ℓ . A general discussion of our approach follows.

We assume as given the topology of the network including SIMP locations. We also assume a given demand matrix $[\gamma_{ij}]$. Let λ_ℓ be the traffic on link ℓ . The (continuous) capacity assignment problem can be formulated as

$$\min_{C_\ell, C_S} T = \frac{1}{\gamma} \left[\sum_{\ell=1}^L \lambda_\ell T_\ell + \lambda_S T_S \right]$$

subject to the constraint that the overall budget B is specified where B is

$$B = \sum_{\ell=1}^L b_\ell C_\ell + b_S C_S$$

and b_ℓ and b_S are cost functions of the capacities C_ℓ and C_S respectively.

In solving the capacity assignment problem, we need to know the total allowable budget and unit cost of ground and satellite channel capacities. We assume the unit cost of ground channel capacities are the same, so we can normalize all costs by a unit cost of ground channel capacity. In obtaining

the following results, we further assume a ratio of 1 to 10 for satellite and ground channel capacity costs, i.e.,

$$\frac{1 \text{ unit cost of ground channel capacity}}{10 \text{ unit cost of satellite channel capacity}} =$$

Our program is sufficiently general so that we can assume any ratio between satellite capacity and ground capacity costs. Perhaps a more realistic ratio to use today is a 1 to 3 proportion.

Without going into the details of the capacity assignment optimization scheme, which is quite similar to the work of Kleinrock [21] and which is described elsewhere [17], let us discuss the computer program flow chart in Figure 3 which combines the routing and capacity assignment algorithms in an overall procedure. The program first obtains a solution for the routing assignments including overall average delay T_R , the routing indices g_{ij} and the channel traffic rates λ_s and λ_ℓ . Then with the λ_s and λ_ℓ from the routing calculations and with a given total budget B and the assumed unit costs of satellite and ground channel capacities b_s and b_ℓ as inputs to the capacity assignment routine, we obtain the capacity assignments C_s and $\{C_\ell\}$ and a new figure for overall average delay T_C . If $T_C \leq T_R$ then the entire procedure is repeated with the previously computed values of C_s , $\{C_\ell\}$ and g_{ij} used as input to the routing subroutine. The process is iterated until $T_C = T_R$ to within a specified accuracy or until the maximum time limit allowed for the optimization procedure is reached.

Using the output data from the routing subroutine run given in the previous section, the iterative cycle is carried out for the network in Figure 2. The input data consists of the original demand matrix, initial

values of $C_g=50$ K bits/sec and $C_s=1500$ K bits/sec and an assumed total budget B in cost units and $b_g=1$ and $b_s=0.1$. The final results for $\{C_g, C_s\}$ and $\{\lambda_g, \lambda_s\}$ after 8 iterations for $B = 1,700,000$ cost units are given in Table 2, and those for $B = 950,000$ cost units after 4 iterations are given in Table 3. For the total budget $B = 1,500,000$ the minimum overall average delay was $T = .037$ sec. whereas for $B = 950,000$, $T = 1.435$ secs. We have performed a number of runs with different values of budget B and the results show that as total allowable budget increases the overall average delay decreases as seen in Figure 3. Also the optimum satellite capacity decreases and the ground channel capacities increase as the total allowable budget increases. This is because we have assumed a 10 to 1 ratio between the cost of unit ground capacity to unit satellite capacity. As the total budget grows we have more money to spend on the relatively more expensive ground channel capacities. As ground channel capacities become large, packets tend to go via ground subnet since the minimum satellite propagation delay is high (0.26 secs). We can show in general that the T vs B curve is a hyperbola and have derived formulas for the asymptote S which are given in a companion paper [17].

THE MASTER SCHEME

In this section we will describe a scheme in which retransmission traffic from a multi-access channel is sent via the ground net. This scheme will be denoted here as the MASTER (Multi-Access Satellite With Terrestrial Retransmission) plan. Under the MASTER plan, the up-link of the satellite channel is a slotted multi-access channel and the down-link is a broadcast channel. Unlike slotted ALOHA, whenever a packet collision occurs in the multi-access

		Optimum Traffic Rates (packets/sec) λ_l	Optimum Capacities (K bits/sec) C_l
Satellite Channel		2.15	15.966
Ground Channel	1	67.95	55.196
	2	120.06	88.238
	3	119.80	88.082
	4	92.43	70.967
	5	159.54	112.292
	6	159.43	112.226
	7	92.46	70.985
	8	120.00	88.201
	9	119.81	88.088
	10	67.80	55.097
	11	67.81	55.109
	12	120.09	88.259
	13	119.67	88.001
	14	92.46	70.984
	15	159.55	112.295
	16	159.33	112.163
	17	92.46	70.987
	18	120.00	88.203
	19	119.71	88.025
	20	67.90	55.165

The minimum delay = .0369 seconds

Table 2

OPTIMUM TRAFFIC RATES AND CHANNEL CAPACITIES FOR B=1,700,000 COST UNITS

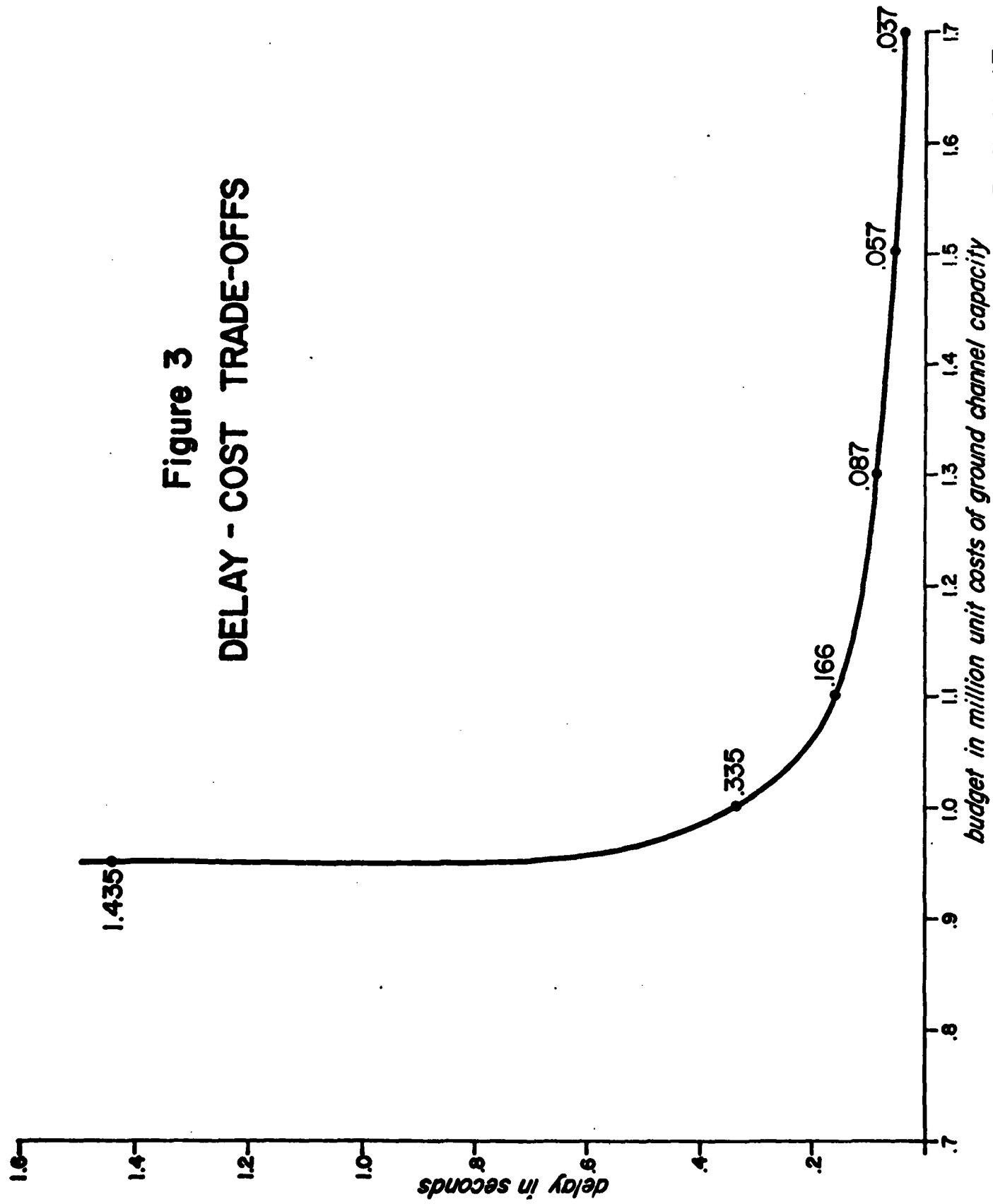
		Optimum Traffic Rates (packets/sec) λ_l	Optimum Capacities (K bits/sec) C_l
Satellite Channel		270.13	511.770
Ground Channel	1	93.61	47.405
	2	79.67	40.391
	3	88.29	44.730
	4	77.50	39.298
	5	81.71	41.417
	6	102.99	52.125
	7	77.61	39.351
	8	82.43	41.780
	9	86.86	44.009
	10	93.62	47.413
	11	91.26	46.222
	12	82.50	41.815
	13	85.94	43.547
	14	79.97	40.541
	15	82.07	41.598
	16	103.11	52.185
	17	79.97	40.539
	18	80.43	40.773
	19	89.34	45.257
	20	91.14	46.166

The minimum delay = 1.435 seconds

Table 3

OPTIMUM TRAFFIC RATES AND CHANNEL CAPACITIES FOR B=950,000 COST UNITS

Figure 3
DELAY - COST TRADE-OFFS



channel, the MASTER scheme does not use the satellite channel for retransmission of the rejected packet. Instead the rejected packet is sent via the ground net from the source SIMP directly to the destination IMP without having to first go to its regional SIMP. By sending retransmitted packets via the ground net, overall average delay can be reduced significantly especially if the traffic on the satellite channel is heavy. Moreover, the MASTER plan eliminates all possibilities of instabilities that might occur in a slotted ALOHA channel under very heavy traffic situations [13]. Thus we see that the MASTER channel has no retransmitted packets. Since on the down-link broadcast channel each SIMP can hear its own transmissions, if a packet collision occurs, the source SIMP could reroute the rejected packet through the ground net without waiting for an acknowledgement from the destination SIMP. Moreover, after sending a packet, a SIMP can immediately send again in the next slot without waiting for the results of previous transmissions.

Now let us turn our attention to a more detailed discussion of the MASTER plan. Consider the satellite channel model depicted in Figure 4. Let λ_{σ} [pkts/sec] be the average input rate from SIMP σ to the satellite channel. The probability q_{σ} that a packet will be transmitted from SIMP σ in a given time slot, or equivalently the traffic rate in a given time slot is:

$$q_{\sigma} = \frac{\lambda_{\sigma}}{\mu_s C_s}$$

where C_s [bits/sec] is satellite channel capacity and $1/\mu_s$ [bits/sec] is the packet length.

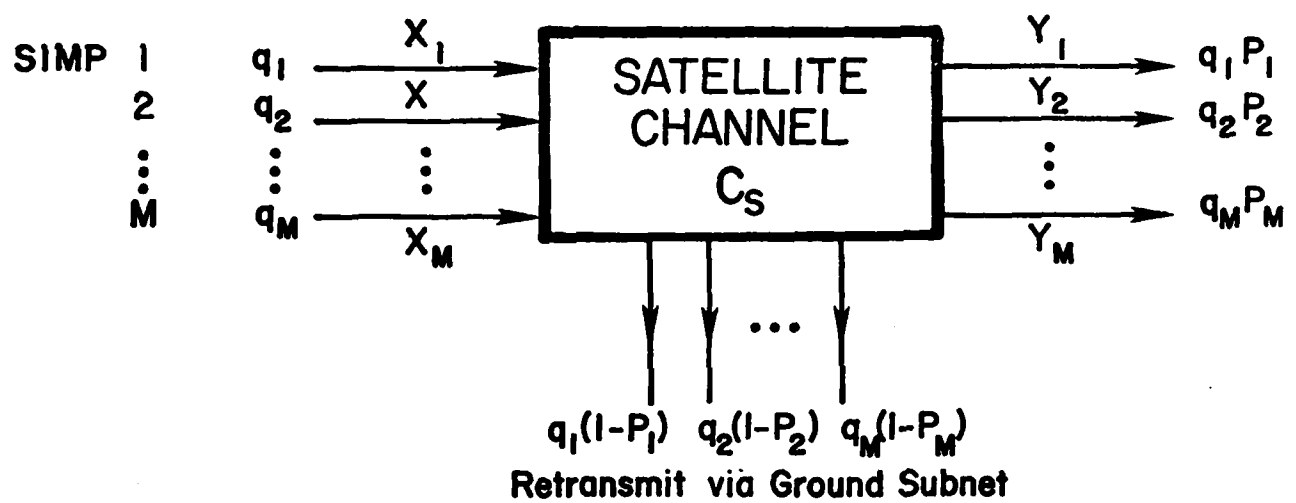


Figure 4

THE SATELLITE SUBNET MODEL UNDER MASTER

We define random sequences $X_\sigma(k)$ and $Y_\sigma(k)$ as

$$X_\sigma(k) = \begin{cases} 1 & \text{if SIMP } \sigma \text{ transmits in time slot } k \\ 0 & \text{otherwise} \end{cases}$$

$$Y_\sigma(k) = \begin{cases} 1 & \text{if SIMP } \sigma \text{ successfully transmits in time slot } k \\ 0 & \text{if otherwise (i.e., no transmission or collision)} \end{cases}$$

Then it follows that

$$\Pr[X_\sigma=1] = q_\sigma$$

By assuming that the streams $X_\sigma(k)$ from different SIMPs are independent we readily obtain

$$\Pr[Y_\sigma=1] = q_\sigma P_\sigma$$

where $P_\sigma = \prod_{\tau \neq \sigma} \bar{q}_\tau$, with $\bar{q}_\tau = 1 - q_\tau$, is the probability of success of a packet from SIMP σ . The normalized throughput S or successful transmission per slot time, of the satellite channel is thus

$$S = \sum_{\sigma} \Pr[Y_\sigma=1]$$

$$= \sum_{\sigma} q_\sigma P_\sigma$$

Note that packets from SIMPs traverse the satellite channel once and only once under this operational scheme. If we assume the satellite channel transmission

errors and nodal processing delay to be small, and buffer storage to be sufficiently large, then the average delay incurred by a packet when going through the channel can be found by using response time formulation of an M/D/1 queueing system plus the satellite propagation delay τ_{\min} of approximately .26 secs. The result of a deterministic service time system is used, since in a slotted channel all packets are of fixed length, any messages with length less than a specified full packet size are filled with blank characters to form a full size packet. Thus the average delay experienced by a packet travelling from SIMP σ through the satellite channel is

$$T_{\sigma} = \tau_{\min} + \frac{1}{\mu_s C_s} \left\{ 1 + \frac{\lambda_{\sigma}}{2(\mu_s C_s - \lambda_{\sigma})} \right\}$$

The overall average packet delay of the satellite channel τ_s [secs/pkt] can be obtained by averaging over all T_{σ} 's

$$\begin{aligned} \tau_s &= \frac{1}{\lambda_s} \sum_{\sigma} \lambda_{\sigma} T_{\sigma} \\ &= \tau_{\min} + \frac{1}{\mu_s C_s} + \frac{1}{2\lambda_s} \sum_{\sigma} \frac{q_{\sigma}^2}{1 - q_{\sigma}} \end{aligned} \quad (7)$$

where

$$\lambda_s = \sum_{\sigma} \lambda_{\sigma}$$

and

$$q_{\sigma} = \frac{\lambda_{\sigma}}{\mu_s C_s}$$

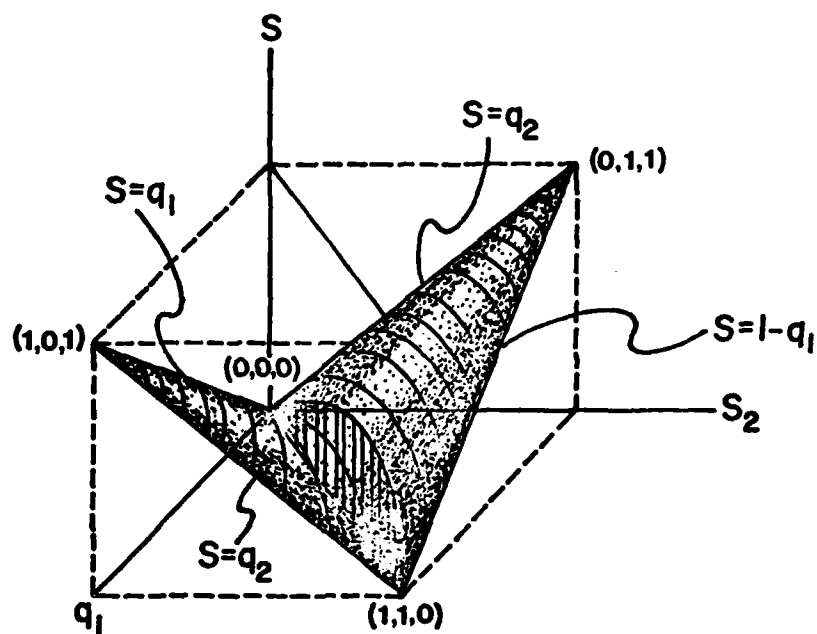


Figure 5
THROUGHPUT SURFACE GIVEN BY EQUATION (8)

It can be shown that for a given total traffic rate $\lambda_s = \sum_{\sigma} \lambda_{\sigma}$ [pkts/sec] τ_s takes on its minimum when $\lambda_{\sigma} = \lambda_s/M$, for all $\sigma=1,2,\dots,M$. We can also show that the condition $\sum_{\sigma} q_{\sigma} = 1$ is necessary to achieve the maximum throughput S . From these conditions we immediately see that the best throughput delay trade-off is attained for $q_{\sigma} = 1/M$, $\sigma=1,2,\dots,M$.

This condition, although the optimum operating point, is difficult to attain, since the traffic rates from SIMPs are most likely not equal. It is difficult to study the trade-off for cases with a large number of heterogeneous SIMPs. We can obtain some understanding of how S varies by investigating the case of $M=2$. For this case, the throughput rate of the satellite channel is

$$S = q_1(1-q_2) + q_2(1-q_1) = q_1 + q_2 - 2q_1q_2 \quad (8)$$

A plot of the throughput S versus the traffic rates q_1 and q_2 is given in Figure 5. In the throughput as shown in Figure 5, the surface takes the form of a saddle with the valley located on the plane $q_1=q_2$ which includes the S -axis and is at a 45° angle from the planes $q_1=0$ and $q_2=0$ (the slashed, diagonal plane shown in Figure 5 and ridge located on the plane $q_1+q_2=1$ which is perpendicular to the plane $(q_1=q_2)$ and passes by upper-corner end points $(1,0,1)$ and $(0,1,1)$ of the throughput surface. Note that the condition

$$q_s = q_1 + q_2 \quad (9)$$

corresponds to a straight line of slope -1 on q_1 - q_2 plane. As stated above if

$$q_1 + q_2 = q_s = 1 \quad (10)$$

we obtain the maximum boundary of the throughput S . The ridge of the saddle surface of the throughput S is thus the maximum boundary. For this simple case, we can further prove that, given the condition in Equation (9), S attains its minimum for

$$q_1 = q_2 \quad (11)$$

which is the line with slope 1 on q_1 - q_2 plane (the line at 45° degree angle with q_1 -axis and q_2 -axis). The valley of the saddle surface of the throughput is thus the minimum boundary. Also, the overall average packet delay, τ_s , of the satellite channel for $m=2$ case is

$$\tau_s = \tau_{\min} + \frac{1}{2\lambda_s} \left(\frac{q_1^2}{1-q_1} + \frac{q_2^2}{1-q_2} \right) + \frac{1}{\mu_s C_s} \quad (12)$$

Again, the delay τ_s is symmetrical with respect to q_1 and q_2 . In Figure (6) we show the level curves of the delay surface in the (q_1, q_2) plane. Given the condition in Equation (9) the delay τ_s reaches its minimum for $q_1=q_2$. The minima are shown by a dashed line intersecting the level curves.

The above results are preliminary. We have not computed numerical examples for the MASTER scheme. These will be given in a companion paper by the authors [17].

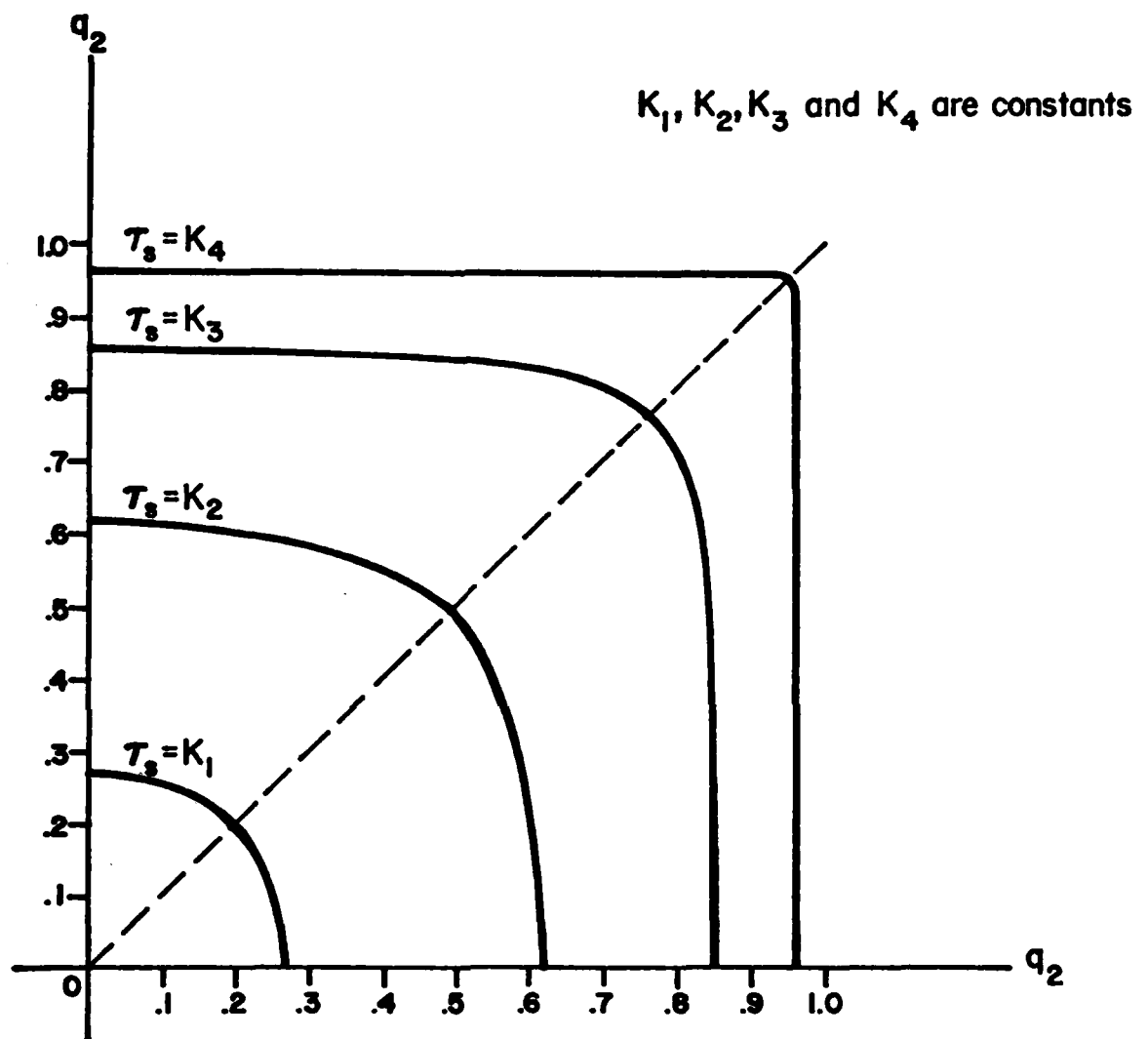


Figure 6
LEVEL CURVES OF DELAY SURFACE
GIVEN BY EQUATION (12)

CONCLUSIONS

In this paper we have presented some of the important design issues for mixed media packet switching networks. Satellite packet switching has considerable promise for low cost, high bandwidth data communications. However there is inherent high delay in satellite links which do not appear in ground links. Therefore a mix of the two communications media seems to offer the best of both worlds. In this paper we have examined a number of trade offs which offer guidelines for the design and optimum utilization of mixed media networks.

We have introduced a new communications scheme called MASTER (Multi-Access Satellite with Terrestrial Retransmission). We believe that the MASTER scheme offers significant advantages over slotted ALOHA especially when the multi-access satellite channel is heavily loaded.

We have not explored the possibility of sending network control information along the ground and using the satellite for bulk data transmission. However the extension is logical and not difficult to analyze. Another logical extension of the MASTER scheme is to use the satellite channel on a reservation basis, such as suggested by Crowther, et al [10], Roberts [11] and Binder [14], but use the ground channel to set up reservations. We plan to explore this idea in a subsequent paper.

REFERENCES

- [1] ROBERTS, LAWRENCE G. and WESSLER, BARRY D.: "The ARPA Network," in *Computer-Communication Networks*, edited by Norman Abramson and Franklin F. Kuo, Prentice-Hall, Inc., 1973, pp. 485-500.
- [2] HEART, F.E.: "The ARPA Network," *Proceedings of the NATO Advanced Study Institute on Computer Communication Networks*, edited by R.L. Grimsdale and F.F. Kuo, Noordhoff International Publishers, 1975, pp. 19-34.
- [3] WESSLER, BARRY D. and HOVEY, RICHARD B.: "Public Packet-Switched Networks," *DATAMATION*, July 1974, pp. 85-87.
- [4] BUTTERFIELD, S.; RETTBERG, R.; WALDEN, D.: "The Satellite IMP for the ARPA Network," *Proceedings of the Seventh Hawaii International Conference on System Sciences - Subconference on Computer Nets*, Western Periodicals Co., January 1974, pp. 70-73.
- [5] ABRAMSON, NORMAN: "The ALOHA System," in *Computer-Communication Networks*, edited by Norman Abramson and Franklin F. Kuo, Prentice-Hall, Inc., 1973, pp. 501-518.
- [6] KUO, FRANKLIN F. and ABRAMSON, NORMAN: "Some Advances in Radio Communications for Computers," *Digest of Papers - COMPCON '73*, San Francisco, February 1973, pp. 57-60.
- [7] BINDER, RICHARD, et al: "ALOHA Packet Broadcasting - A Retrospect," *AFIPS National Computer Conference Proceedings*, Vol. 44, 1975, pp. 203-215.
- [8] ROBERTS, LAWRENCE G.: "ALOHA Packet System With and Without Slots and Capture," *ARPANET Satellite System Note #8 (NIC 11290)*, Stanford Research Institute, June 1972.
- [9] ABRAMSON, NORMAN: "Packet Switching with Satellites," *ALOHA SYSTEM Technical Report B73-2*, University of Hawaii, March 1973; also published in *AFIPS National Computer Conference Proceedings*, Vol. 42, New York, June 1973, pp. 695-702.
- [10] CROWTHER, W., et al: "A System for Broadcast Communications: Reservation ALOHA," *Proceedings of the Sixth Hawaii International Conference on System Sciences*, Western Periodicals Co., January 1973, pp. 371-374.
- [11] ROBERTS, LAWRENCE G.: "Dynamic Allocation of Satellite Capacity Through Packet Reservation," *AFIPS National Computer Conference Proceedings*, Vol. 42, June 1973, p. 711.
- [12] KLEINROCK, LEONARD and LAM, SIMON S.: "Packet-Switching in a Slotted Satellite Channel," *AFIPS National Computer Conference Proceedings*, Vol. 42, June 1973, p. 703.

- [13] LAM, SIMON S.: "Packet-Switching in a Multi-Access Broadcast Channel with Application to Satellite Communication in a Computer Network," Ph. D. Thesis, UCLA; also available as *UCLA-ENG-7429*, UCLA Computer Science Department, April 1974.
- [14] BINDER, RICHARD: "A Dynamic Packet Switching System for Satellite Broadcast Channels," *ALOHA System Technical Report B74-5*, University of Hawaii, August 1974.
- [15] DAVIES, D.W.: "A Review of Computer Communications Technology," *Proceedings of the NATO Advanced Study Institute on Computer Communication Networks*, edited by R.L. Grimsdale and F.F. Kuo, Noordhoff International Publishers, 1975, pp. 1-17.
- [16] FRANK, H.; FRISCH, I.T. and CHOU, W.: "Topological Considerations in the Design of the ARPA Computer Network," *AFIPS Spring Joint Computer Conference Proceedings*, Vol. 36, 1970, pp. 543-549.
- [17] HUYNH, D.; KOBAYASHI, H. and KUO, F.F.: "Optimal Design of Mixed Media Packet Switching Networks: Routing and Capacity Assignment" to be published in *IEEE Trans. on Communication*, Autumn 1976.
- [18] CROWTHER, W.R.; HEART, F.E.; McKENZIE, A.A.; McQUILLAN, J.M.; and WALDEN, D.C.: "Issues in Packet Switching Network Design," *AFIPS National Computer Conference Proceedings*, Vol. 44, 1975, pp. 161-175.
- [19] FULTZ, G.L.: "Adaptive Routing Techniques for Message Switching Computer-Communications Networks," Ph. D. Thesis, UCLA; also available as *UCLA-ENG-7252*, UCLA Computer Science Department, July 1972.
- [20] BOX, M.J.: "A New Method of Constrained Optimization and a Comparison with Other Methods," *The Computer Journal*, August 1965, p. 42.
- [21] KLEINROCK, L.: *Communication Nets: Stochastic Message Flow and Delay*, Dover Publications, New York, 1964.
- [22] JACKSON, J.R.: "Job Shop-like Queueing Systems," *Management Science*, Vol. 10, No. 1, October 1963, p. 131.

2-8

DT